DETECTION OF CANCER CELL PREDICTION IN BLOOD SAMPLES USING WATERSHE ALGORITHM

A.Selvarani , **A.Mathangopi M.E.,** Associate Professor
Department of CSE
Sri Muthukumaran Institute Of Technology.

Abstract: *Acute leukemia is a disease of the leukocytes and their precursors. It is characterized by the appearance of immature, abnormal cells in the bone marrow and peripheral blood and frequently in the liver, spleen, lymph nodes, and other parenchymatous organs. The paper presents the preprocessing methods of the leukemic blast cells image in order to generate the features well characterizing different types of cells. The solved problems include: the segmentation of the bone marrow aspirate by applying the watershed transformation, selection of individual cells, and feature generation on the basis of texture, statistical and geometrical analysis of the cells.*

1. INTRODUCTION

Leukemia is a cancer that begins in the bone marrow. It is caused by an excessive production of immature leucocytes that replace normal blood cells (leukocytes, red blood cells, and platelets). It causes the body to be exposed to many diseases with no possibility to fight them for lack of defense.

Without treatment, this cancer is the cause of many deaths. Based on the statistics it is been realized that it is the fifth and sixth cause of death among men and women with this cancer.

Leukemia is curable if it is detected and treated in at early stage. Is detection starts with a complete blood could. If there are abnormalities in this count, a study of morphological bone marrow smear analysis is done to confirm the presence of leukemic cells. In this study, a pathologist observes some cells under a light microscopy looking for abnormalities presented in the nucleus or cytoplasm of the cells in order to classify the abnormal cells in their particular types and subtypes of leukemia.

This classification is very important as it determines which treatment is given. This study has an error rate between 30% and 40% depending on the pathologist experience and the difficulty to distinguish leukemia types and subtypes. A flow cytometry test is highly accurate to classify leukemias but it is very expensive and not all the hospitals have the equipment to perform it. The classification of leukemia types and subtypes facilitate the physicians' work in deciding what treatment is the best or a given cell type (lymphoctic or myelogenous) and disease progress (acute or chronic). This paper presents a preprocessing method of the leukemic blast cells image in order to generate the features well characterizing different types of cells. The recognition of the blast cells in the bone marrow of the patients suffering from myelogenous leukemia is a very important step in the recognition of the development stage of the illness and proper treatment of the patients [2,3,9]. The percentage of blasts is a major factor at defining various subtypes of acute myeloid leukemia. According to French-American- British (FAB) standard, 8 acute leukemia types are classified on the basis of the ratio of myelo/monoblasts, the number of erythroid precursors or non-erythroid cells as well as megacarioblasts cells. It is known that proper treatment of leukemia requires not only recognition of different stages of the development of the blasts but also calculation of their quantity in the aspirated bone marrow.

II TYPES OF CELLS

There exist many different cell types in the bone marrow. The most known and recognized abnormal cells include monoblasts, promonocytes, monocytes, myeloblasts, promyelocytes, myelocytes, metamyelocytes, proerythroblasts, basophilic erythroblasts, polychromatic erythroblasts, orthochromatic erythroblasts, lymphocytes, plasmocytes, megacaryoblasts, megacaryocytes, etc [2,3]. The variety of cells occurring in the bone marrow demands a high expertise of the analyst, which is usually verbal one. For improving the reliability of the analysis and diagnosis, computer based digital image processing offers a useful tool. This paper is dedicated to the task of feature generation for the automatic blast cell recognition. The well-defined features should suppress the differences among the cells belonging to the same class and amplify them for cells belonging to different classes.

The presented solution may be treated as the first step in building up an automatic system able to recognize different blood cells.

III MEDICAL IMAGE PROCESSING

A) Medical Imaging

In the clinical context, medical imaging is generally equated to radiology or "clinical imaging" and the medical practitioner responsible for interpreting (and sometimes acquiring) the images is a radiologist. Diagnostic radiography designates the technical aspects of medical imaging and in particular the acquisition of medical images. The radiographer or radiologic technologist is usually responsible for acquiring medical images of diagnostic quality, although some radiological interventions are performed by radiologists. While radiology is an evaluation of anatomy, nuclear medicine provides functional assessment.

As a field of scientific investigation, medical imaging constitutes a sub-discipline of biomedical engineering, medical physics or medicine depending on the context: Research and development in the area of instrumentation, image acquisition (e.g. radiography), modeling and quantification are usually the preserve of biomedical engineering, medical physics and computer science; Research into the application and interpretation of medical images is usually the preserve of radiology and the medical sub-discipline relevant to medical condition or area of medical science (neuroscience, cardiology, psychiatry, psychology, etc.) under investigation. Many of the techniques developed for medical imaging also have scientific and industrial applications.

B) Image Segmentation

Segmenting the nucleus and cytoplasm of leukocytes from bone marrow images is a very difficult task, as the images show heterogeneous staining and high-cell population as shown in figure below.

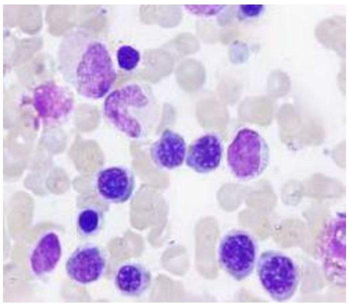

Figure 1. The exemplary image of the bone marrow smear of the acute leukemia patient containing different blast cells.

Some segmentation techniques such as thresholding, edge detection, pixel clustering, and growing regions have been combined to extract the nucleus and cytoplasm of leukocytes. These techniques could be applied as the images showed uniform backgrounds and high contrast that appropriately defined the objects of interest. In this paper an approach is been proposed in which a segmentation algorithm based on color and texture pixels features is performed that can work in bone narrow images showing heterogeneous staining.

IV PROPOSED METHODS

The task of segmentation of the image is focused on the automatic recognition and separation of each cell for further processing, in order to obtain stable features, useful in recognition of the cell.

A) Cell Segmentation

Three important steps is involved in segmenting the image given as:

1. Segmentation of cellular elements,
2. Identification of nucleus and cytoplasm, and
3. Separation of overlapped blood cells.

Morphological operations are used in solving the segmentation. The morphological operations aim at extracting relevant structures of the image by probing the image with another set of a known shape called structuring element, chosen as the result of prior knowledge concerning the geometry of the relevant and irrelevant image structures. The most known morphological operations include erosion, dilation, opening and closing [6]. The morphological approach to image segmentation combines regions growing and edge detection techniques.

It groups the pixels around the regional minima of the image. The boundaries of adjacent grouping are precisely located along the crest lines of the gradient image. In our experiments, watershed transformation operation is performed to obtain the result. The watershed transformation [4,6] takes its origin from the topographic interpretation of the gray scale image. According to the law of gravitation, the water dropped on such surface will flow down until reaches a minimum. The whole set of points of the surface, whose steepest slope paths reach a given minimum, constitutes the catchment's basis associated with this minimum. The watersheds are the zones dividing adjacent catchment's basins. In numerical implementation of the watershed algorithm the original image is transformed so, as to output an image whose minima mark relevant image objects and whose crest lines correspond to image object boundaries. In this way the image is partitioned into meaningful regions that may correspond for example to the individual blast cells. In our experiments, Watershed algorithm is used for implementation using Matlab platform [10]. The applied procedure of the image segmentation and cell separation consists of the following stages:

- Transformation of the original image into gray scale.
- Transformation of the gray image to binary one by applying the biased segmentation.
- Application of closing and erosion operations to smooth the contours and to eliminate the distortions.
- Generation of the map of distances from the black pixel to the nearest white pixels.
- Application of the watershed algorithm for the image segmentation.

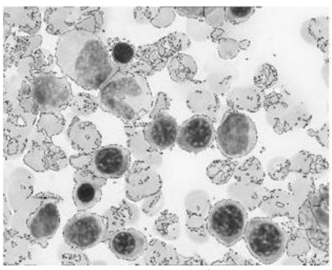

Figure 2. The segmented image of the bone marrow smear.

B) Features Extraction.

Applying decomposition model on image, as a texture analysis is used, to separate texture into its structural and stochastic components. The blood cells images present heterogeneous textures, hence both periodical and random textures can be found in such images. Additional motivations for choosing this model were its similarity relation with the human visual perception system, and its invariant properties to translation, rotation, and scale.

The wold decomposition model interprets the image texture by means of the sum of three mutually orthogonal components, a harmonic field, a generalized evanescent field, and a stochastic field. The perceptual characteristics of these fields can as periodicity, directionality, and randomness, respectively, according to the three most important human perception dimensions.

C) The geometrical features

The important information is contained in the geometrical shapes and parameters [9] associated with them. Various cells differ greatly with the size. For example the orthochromatic erythroblasts have the size of 8-10 micrometer, while megakcariocyte may be up to 100 micrometer. The shapes of different blasts are either round, oval or kidney-shaped. Following geometrical features of the cells is been considered.

- Radius –measured by averaging the length of the radial line segments defined by the centroid and border points
- Perimeter - the total distance between consecutive points of the border
- The ratio of the perimeter and radius
- Area – the number of pixels on the interior of the cell, defined separately for the nuclei and for the whole cell; as the features we assume the area of the nucleus and the ratio of the areas of the nucleus and the whole cell
- The area of convex part of the nucleus
- Compactness – given by the formula: perimeter2/area
- Concavity – the severity of concavities in a cell
- Concavity points – the number of concavities, irrespective of their amplitudes
- Symmetry – the difference between lines perpendicular to the major axis to the cell boundary in both directions
- Major and minor axis lengths.

D) Statistical Features Process

The next set of features has been generated on the basis of the intensity distribution of the image. The histograms and gradient matrices of such intensity have been determined for three color components R, G and B. On the basis of this the following features have been generated: the mean value and variance of the histogram and the gradient matrix of the image of the nucleus and the whole cell (24 features), skew ness and kurtosis of the histogram and gradient matrix of the whole cell (12 features). All these features have been calculated for three colors. 36 features have been generated in this way. All numerical experiments of feature generation have been implemented on the platform of Matlab [10].

E) Identificattion Process

From the regions obtained in the segmentation process, by analyzing their shape, color, and special relation with respect to other regions to determine whether and analyzed region is a nucleus or a leukocyte.

The features that were used to recognize cellular elements are: circularity to measure the perimeter complexity of a circular object (circularity = perimeter2 / (4πarea)), eccentricity to find out how much the object deviates from being circular (eccentricity = dist(center, focus)), color to determine if a region is darker than other, and containment proportion to establish whether a region contains or is contained by another region.

F) Classification Process

The suitable recognition of leukemia cells requires the definition of good descriptive features that facilitate their classification. In this phase geometric feature is been extracted, and statistical, texture, and size ratio features from regions obtained in the segmentation process (nucleus, cytoplasm, and whole cell) and analyzed these features to identify types and subtypes of acute leukemia.

V RESULTS OBSERVATIONS

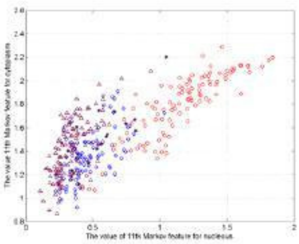

Figure 3. The distribution of cell locations in plane formed by 11th Markov feature corresponding to cytoplasm and nucleus.

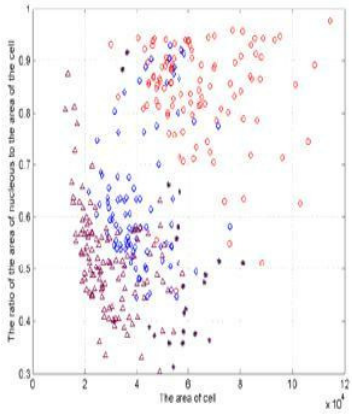

Figure 4. The distribution of cell locations in the plane formed by 2 geometrical features.

VI CONCLUSIONS

The paper has presented the image processing approach to the recognition and classification of the leukemia cells. The most important points of this approach are: segmentation of the image of bone marrow aspirate using watershed algorithm, the extraction of individual cells from the image, automatic generation of different features of the cell, assessment of the feature quality of the cells using analysis of distribution, correlation and principal component analysis, application of support vector machine for final recognition and classification of cells.

VII REFERENCES

[1] S. Haykin, "Neural networks, comprehensive foundation", Prentice Hall, New Jersey, 1999 [2] H. Hengen, S. Spoor, M. Pandit, "Analysis of blood & bone marrow smears, SPIE Med. Imag., San Diego, 2002 [3] K. Lewandowski, A. Hellmann, "Haematology atlas", Multimedia Medical Publisher, Gdansk, 2001

[4] O. Lezoray, H. Cardot, "Cooperation of color pixel classification schemes and color watershed", IEEE Trans. Image Processing, vol. 11, pp. 783-789, 2002 [5] O. L. Mangasarian, P. Lagrangian, "Support Vector Machines", Journal of Machine Learning , 161-177, 2001

[6] P. Soile, "Morphological image analysis, principles and applications", Springer, Berlin, 2003

[7] V. Vapnik, "Statistical Learning Theory", Wiley, N.Y., 1998 .

[8] T. Wagner, "Texture analysis" (in Jahne, B., Haussecker, H., and Geisser P., (Eds.), Handbook of Computer Vision and Application), Academic Press, pp. 275-309, 1999 .

[9] W. Wolberg, W. N. Street, O. L. Mangasarian, "Machine learning to diagnose breast cancer from image-processed features", Rep. of Uni. Wisconsin, 1994 .

[10]Matlab user manual – Image processing toolbox, MathWorks, Natick, 1999.